Assessing the Impact of Adding Scenario-based Augmented Reality Training to Traditional Textbook Training on Accurately Identifying Airway Obstruction Symptoms

Research Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in the undergraduate colleges of The Ohio State University

by

Lauren A. Mansour

The Ohio State University
April 2020

Research Thesis Committee:
Project Advisor: Dr. Emily S. Patterson, College of Medicine, School of Health and Rehabilitation Sciences
Committee Member: Dr. Megan Gregory, College of Medicine
Honors Program Director: Dr. Melissa Quinn, Division of Anatomy, College of Medicine

Acknowledgements

I would like to express my sincerest gratitude to my advisor and Principal Investigator, Dr. Emily Patterson, for her invaluable guidance and support throughout this project. I would also like to thank my research teammates Olivia Hernandez, Varshita Chirumamilla, and Sabrina Amin for their hard work and collaboration. I would like to thank Dr. Christopher San Miguel, Dr. Michael Barrie, and TJ Seelig for their constant willingness to support this project and their invaluable medical knowledge. In addition, I would like to thank Christen Sushereba and Laura Militello from Unveil, LLC, for their support in analyzing the data and appreciating the significance of the project. I would also like to thank Unveil LLC as a whole for allowing me to use their technology to complete this project. Lastly, I would like to thank my parents and twin sister for their unfailing support and encouragement throughout the entire project.

This work is supported by the US Army Medical Research and Materiel Command under Contract No. W81XWH-17-C-0162. The views, opinions and/or findings contained in this paper are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

Table of Contents

EXECUTIVE SUMMARY

Recent developments in the portability, flexibility, and affordability of augmented reality (AR) technology allow it to address some limitations of current medical training. This pilot study aims to determine whether AR technology can enhance the detection and identification of subtle cues used to make accurate medical diagnoses. We randomly divided five participants into a control group, who received diagnostic training by reading a standard textbook, and an experimental group, who received diagnostic training by a combination of reading a standard textbook and AR-based training. Participants then encountered a tension pneumothorax scenario on the AR patient and were asked to list the symptoms they observed before making a final diagnosis. We recorded both the total amount of time participants required to make a final diagnosis as well as the number of correctly identified symptoms, and the results of the control and experimental groups were compared with a t-test. We found no significant difference between groups in the time needed to make a diagnosis. However, the AR-trained participants correctly identified a significantly greater number of symptoms (average: 12 symptoms) compared to the textbook-trained participants (average: 5 symptoms). These results indicate the potential of AR-based training to enhance detection and identification of cues in medical diagnostics.

INTRODUCTION

Immersive technologies, including augmented, virtual, and mixed reality are becoming increasingly accessible to medical educators. Recent developments in the portability, flexibility, and affordability of augmented reality (AR) technology make it a tool that can address the limitations of current medical training. AR differs from other immersive technologies, such as virtual reality (VR), in that the visuals designed in AR are superimposed over the tangible - or real-world environment. With VR technology, the user can see and interact only in the virtual environment (cf., Kellen et al., 2019), but AR technology allows the user to see and interact with a visual overlay superimposed on tangible objects in the physical environment. AR is more beneficial for training purposes because users can learn to not only recognize critical cues, but to also practice manual skills, thereby improving perception-action cycles, or the continuous interplay between thought and interaction with surroundings. This gives AR a distinct and crucial advantage over VR, which does not allow users to interact with the physical environment. The clear benefits of AR over VR technology in a medical instruction setting, coupled with emerging advancements in AR, position AR-based training as an encouraging potential method of medical training; however, design principles for effective AR training have not yet been demonstrated in a research setting.

As part of a multi-phase study, this experiment is focused on supporting simulation-based medical education through the use of AR technology. **Figure 1** is an example of an augmented reality patient that a user may see in this pilot study. This figure demonstrates jugular vein distension via a visual overlay on top of a physical manikin.

Figure 1. Image of MART-CM™ Augmented Reality technology from Unveil LLC included with permission, also available at https://www.youtube.com/watch?v=HjVxyiKczME&feature=youtu.be

We aim to identify how this technology can enhance the detection and identification of subtle visual cues from the real world in the process of making accurate medical diagnoses. The improvement in the identification of minute cues from the real world into an AR environment is an example of the macrocognition function of sensemaking, which emphasizes recognizing perceptual cues in a scene and linking those cues with the best explanation or diagnosis (Patterson & Hoffman, 2012). Therefore, our specific research question is: *Is augmented reality training more effective in supporting the identification of critical symptoms and diagnosis of airway obstruction patients than traditional textbook training?* To answer that question, this study is an experiment designed to examine the performance differences between textbook-trained and AR-trained participants in real-life scenarios.

METHODS

Data Collection

Five participants were recruited for this pilot study in order to determine how many study participants would be required to detect a statistically significant difference using a primary measure of time. All pilot study participants were undergraduate students at The Ohio State University majoring in Biomedical Science in the College of Medicine. A recruitment email was sent describing details of participation and informed consent was obtained. The email stated that the experiment would take place in the Leverage Point Engineering Laboratory in Atwell Hall and would last no longer than an hour. These participants were randomly assigned either the standard textbook training or the AR-based training. The order of the trials was determined randomly with a randomization scheme using =RAND in MS Excel™. After completion of each training session, we asked the following five semi-structured interview questions to evaluate perceptions about the experience:

1. You diagnosed the patient with __________. Can you explain your reasoning for that diagnosis?

2. Were there any signs or symptoms that you wish were present in the augmented reality technology?

3. Were there any signs or symptoms that were present in the augmented reality technology that you felt were misleading?

4. **For those in the experimental group**: Did you feel the augmented reality training included: too much, not enough, or just the right amount of information to lead you to a diagnosis?

5. Do you have any other comments about your experience with the technology?

During the standard textbook training, participants were asked to read a textbook page (**Figure 2**) describing an airway obstruction which then provided details on two specific examples of airway obstruction: tension pneumothorax and superheated airway. This information was based upon adaptations of three medical resources (www.trauma.org, www.uptodate.com, and www.nbi.nim.nih.gov). A combat medic consultant reviewed these materials and provided suggestions for modification, but overall, generally validated that the information was accurate. With this condition, the participants had no exposure to the augmented reality system until their post-training evaluation using the technology. Before using the technology for the assessment, a demonstration was given using a hemorrhage case. The script for this example case can be seen in **Figure 3.**

Airway obstruction is a blockage in any part of the airway. The airway is a complex system of tubes that conveys inhaled air from your nose and mouth into your lungs. An obstruction may partially or totally prevent air from getting into your lungs. Some airway obstructions are minor, while others are life-threatening emergencies that require immediate medical attention.

Some of the many causes of an airway obstruction include: inhaling or swallowing a foreign object, allergic reaction, trauma to the airway from an external (gunshot wound to neck) or internal superheated airway) injury accident, breathing in large amounts of smoke from a fire, viral infections, asthma, cystic fibrosis, and many more.

An airway obstruction can be acute, sub-acute, or chronic. An acute airway obstruction means the blockage began suddenly and is worsening rapidly. A sub-acute airway obstruction means airflow blockage is mild and symptoms began recently. The diagnosis of airway obstruction can frequently be made by history and physical examination. To diagnose an acute airway obstruction, look out for signs of respiratory distress including the following: tachypnea (excessively rapid respiration), tachycardia (rapid heart rate), stridor (high-pitched wheezing), hypoxia (low blood oxygen), impairment of oxygenation and ventilation which can lead to cyanosis (bluish discoloration). If left untreated, an acute airway obstruction can lead to respiratory or cardiac arrest. To diagnose a sub-acute airway obstruction, look out for the following symptoms: dyspnea (labored breathing), cough, hemoptysis (coughing up blood), and stridor.

With airway obstruction, it is important to start the intervention quickly, such as intubating the patient, in order to secure the airway.

An example of an airway obstruction is a tension pneumothorax. Tension pneumothorax is the progressive build-up of air within the pleural space, usually due to a lung laceration which allows air to escape into the pleural space, but not to return.

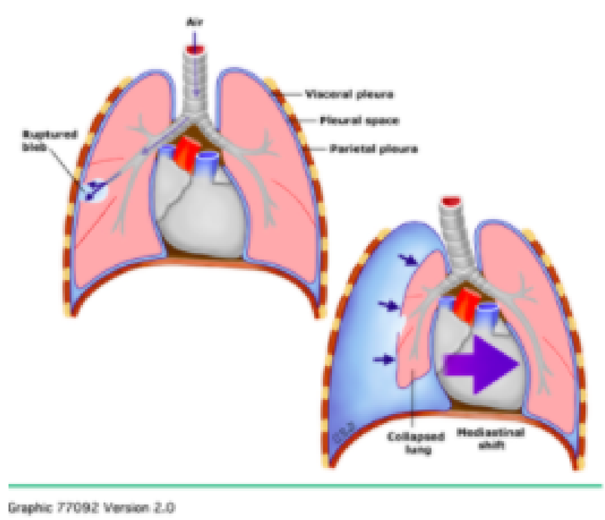

Figure 2. Page 1 of Traditional Textbook

mediastinum (central compartment of the thoracic cavity) to the opposite hemithorax, and obstructs venous return to the heart. This leads to circulatory instability and may result in cardiac arrest. The classic signs of a tension pneumothorax are deviation of the trachea away from the side with the tension, an increased percussion note, a hyper-expanded chest that moves little with respiration, and decreased breath sounds on the affected side. The central venous pressure is usually raised, but will be normal or low in hypovolemic (decreased volume of circulating blood in the body) states.

However these classic signs are usually absent and more commonly the patient is tachycardic (rapid heart rate) and tachypneic (excessively rapid respiration), and may be hypoxic (low blood oxygen). These signs are followed by cardiac arrest with hypotension and subsequent traumatic arrest (no pulses or spontaneous respiratory activity). Breath sounds and hyperresonant percussion notes may be difficult to appreciate and misleading.

The traditional treatment with a tension pneumothorax is to put in a chest tube and treat with a needle decompression with an angiocatheter.

Another cause of an airway obstruction could be a superheated airway. Inhaling superheated air can cause an airway burn which can lead to rapid swelling of the burned tissue in the airway and can quickly block the flow of air into the lungs.

Symptoms of a superheated airway include a cough with increased sputum production, stridor, and dyspnea with rhonchi or wheezing, accessory muscle usage, tachypnea, cyanosis, and stridor.

The traditional treatment for a superheated airway is to intubate the patient.

Information modified from:
http://www.trauma.org/archive/thoracic/CHESTtension.html

https://www.uptodate.com/contents/clinical-presentation-diagnostic-evaluation-and-management-of-central-airway-obstruction-in-adults?search=airway%20obstruction&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1

https://www.ncbi.nlm.nih.gov/books/NBK513261/

Figure 2. Page 2 of Traditional Textbook

I will now begin to lead you through an example answer. When looking at this patient, the first thing I notice is that he is lying in a supine position. He is bleeding out of his right leg and the blood appears to be bright red. He also appears to be tachypneic, which means he is breathing very fast. When looking at his vital signs, I notice that he has low blood oxygen saturation and he has a high heart rate. When looking at his face, it seems he is beginning to be cyanotic; his lips are blue. He also appears to be very pale. Lastly, he looks hypoxic.

Progress to Stage 4

Now, his eyes are closing, and he is even more pale. He is definitely cyanotic now. He appears to be disoriented or confused and he's unresponsive. The first thing I would do is tie a tourniquet around his leg to stunt the bleeding. After looking at all these signs I'm going to make my diagnosis. I believe this patient has a hemorrhage originating in the right leg.

Figure 3. Script of Example Hemorrhage Case

The AR-based training was a presentation that discussed diagnosing patients with airway or breathing complications, including the order of assessment for a trauma patient with breathing complications. In this mnemonic "ABC" assessment, "A" indicates airway, "B" indicates breathing, and "C" indicates circulation. The presentation provided additional details about each of these assessment areas and pathophysiological complications that occur in each one. The presentation then shifted to defining medical terminology relating to airway obstructions, particularly tension pneumothorax and superheated airway. It also guided participants to shift their gaze to look at the appropriate region of the body relating to each symptom. The majority of the presentation material was generated for a related study by a collaboration of three Emergency Physician medical educators at The Ohio State University.

At the conclusion of either the presentation or the textbook reading, participants were invited to approach the table on which the AR patient had been superimposed. Before receiving the post-training test involving the AR patient, participants acclimated to the technology by practicing with an unrelated scenario: hemothorax. There were no restrictions on where participants were permitted to walk, what they could touch, or how they could move in assessing the AR patient. A demonstration of an example answer was then given using a hemorrhage case. This example listed noticeable cues, stated a treatment plan, and then declared a specific diagnosis. This slide-based and AR-training was in addition to the textbook training (**Figure 1**).

The post-training test consisted of analyzing an AR patient suffering from tension pneumothorax. The participant was asked to identify the patient's signs and symptoms, and propose a diagnosis based on these observations. The participant was recorded both with an audio recorder as well as with hand-written notes. The session concluded with a debrief questionnaire requesting comments on both the training method and the overall experience with the technology.

The AR technology that was used in this pilot study was created by Unveil, LLC. It provides various scenarios, which each contain five stages. The first of these five, Stage 1, is the beginning of a trauma or ailment when the patient is relatively stable. Stages 2 - 4 demonstrate the patient's condition progressively declining (e.g., worsening pallor, eyes closing), and in Stage 5, the patient is deceased. The proctor has control over which stage the participant is assessing. The technology also includes gaze-tracking capability, allowing the proctor to evaluate which critical cues are being assessed and which are being missed by tracking participants' eye movements. In this study, Sages 3 and 4 were used, each presenting the same basic symptoms but with differing severity (Stage 4 more severe). The participant was allowed 120 seconds on Stage 3 and then as much time as needed on Stage 4.

Data Analysis

A rubric was used to identify correct and incorrect detection of critical cues in the two stages used in this study. This rubric was developed for this and related studies and was designed in cooperation with an emergency department physician (**Tables 1-4**). One point was allocated for each correct symptom identified, while one point was subtracted for each incorrect symptom identified. Points were not allocated for correct observations given that were not directly related to tension pneumothorax. Diagnoses were ranked on a scale of 0-1 based on accuracy. Tension pneumothorax earned 1 point, pneumothorax earned 0.5 points, and an incorrect diagnosis received 0 points. All other partially correct diagnoses were graded based on their accuracy determined by an emergency department physician. The points earned for symptoms and the diagnosis grades were averaged. The time to diagnosis was averaged for the control group and experimental group and results were compared with a two-tailed t-test.

Table 1: Symptom Recognition Rubric

Category	Symptom	Acceptable statements
Airway status	Airway compromised	Airway not intact
		Airway compromised
Respiratory status	Dyspneic	Dyspneic
		Dyspnea
	Tachypneic	Respiratory rate high
		Respiratory rate more than 22
		Respiratory rate fast
		Tachypneic
		Tachypnea
		Elevated respiratory rate
		Breathing fast
		Quick breathing
		Frequent breaths
		Rapid breathing
		Fast breathing
	Difficulty of respiration	Labored/shallow breathing
		Shortness of breath
		Short of breath

	Abnormal chest movement	Asymmetric rise and fall of chest
	Respiratory sounds	Diminished
		Severely diminished
		Absent
		One side sounds worse than the other
	Hypoxic	Hypoxic
		Hypoxia
		O2 saturation abnormal
		O2 saturation below 95
		O2 saturation below 90 (pathologic level)
		O2 saturation low
	Cyanotic	Cyanotic
		Perioral cyanosis
		Blue lips
	Abnormal vital sign for respiration	O2 saturation below 95
		O2 saturation below 90 (pathologic level)
		O2 saturation low
	JVD	JVD
		Jugular Venous Distension

	Tracheal Deviation	Deviation away from affected side
		Tracheal deviation
Ability to respond	Non-responsive	Eyes closing
		Eyes closed
Poor Skin Signs	Pale	Pale
		Pallor
		White
		Changing skin tone
		Getting progressively paler
Abnormal cardiac vital signs	Tachycardic	Tachycardic
		High heart rate
	Bradycardic	Bradycardic
		Low heart rate

Table 2. Diagnostic Accuracy Rubric

Diagnosis	Credit
Tension pneumothorax	Full credit (1 point)
Pneumothorax	Partial credit (0.5 points)
Anything else	No credit (0 points)

Table 3. Appropriateness of Treatment Rubric

Diagnosis	Credit
Needle decompression	Full credit (1 point)
Chest tube, finger thoracostomy	Partial credit (0.5 points)
Anything else	No credit (0 points)

Table 4. Appropriateness of Diagnostic Studies (after intervention) Rubric

Diagnosis	Credit
Chest x-ray or ultrasound	Full credit (1 point)
CT scan (requires being stable to be sent)	Partial credit (0.5 points)
Anything else	No credit (0 points)

DATA FINDINGS

The primary data point was the time taken to diagnose the AR patient. **Table 5** shows the time it took each participant to diagnose, the diagnosis proposed, and accuracy of the diagnosis (scale 0 - 1). None of the participants diagnosed the patient specifically with tension pneumothorax, but two of the participants in the experimental group diagnosed the AR-patient with pneumothorax, for which half credit was allotted.

Table 5. Diagnostic Accuracy

Participant	Condition	Dx time	Dx timestamp	Diagnosis	Diagnostic Accuracy (out of 1)
1	Experimental	227	3:47	Pneumothorax	0.5
2	Control	155	2:35	Puncturing of right lung causing it to collapse	0.4
3	Experimental	140	2:20	Hemorrhage from the cut bleeding into mediastinum causing pressure on the lungs inhibiting breathing	0
4	Control	126	2:06	Acute respiratory distress	0.2
5	Experimental	175	2:55	Pneumothorax	0.5

Table 6 compares time to diagnosis for the control group compared to the experimental group. The average for the control group was 140.5 seconds with a standard deviation of 20.5 seconds and the average for the experimental group was 180.67 seconds with a standard deviation of 43.78 seconds. The p-value for this measurement is 0.26, which shows that there was no significant difference in the time it took to make a diagnosis between the two groups.

Table 6. Time to Diagnose

	Control	Experimental
	155	227
	126	140
Time to diagnose (s)		175
Average	140.50	180.67
Standard Deviation	20.51	43.78
P-value		0.26

The number of symptoms correctly identified in each stage were counted and the total amount of correct symptoms stated were summed for each participan**t (Table 7).**

Table 7. Number of Accurate Symptoms Observed per Stage

Participant	Condition	S3 symptoms	S4 symptoms	Sum of both stages
1	Experimental	6	8	14
2	Control	2	2	4
3	Experimental	2	7	9
4	Control	4	2	6
5	Experimental	5	8	13

The number of total accurate symptoms observed were averaged in both the control group and experimental group. It was found that an average of 5 symptoms were accurately identified in the control group with a standard deviation of 1.41. For the experimental group, there were an average of 12 symptoms correctly identified with a standard deviation of 2.65. A 2-tailed T-test gave a p-value of 0.032 meaning there was a significant difference in the number of symptoms correctly identified. This is shown in **Table 8.**

Table 8. Number of Accurate Symptoms Observed

	Control	Experimental
	4	14
	6	9
Sum of both stages		13
Average	5	12
Standard deviation	1.41	2.65
P-value		0.032

Table 9 shows the symptoms the participants identified along with the time at which those symptoms were observed. These were then compared to the rubric (**Table 1**) to determine how accurate and relevant to the specific case they were, and points awarded for each stated symptom. One point was awarded for each correct symptom given, one point was deducted for each incorrect symptom given (e.g., jaundice), and no points were awarded for symptoms that were shown in the technology but were not directly related to a tension pneumothorax (e.g. blood from laceration is bright red).

Table 9a. Findings for participant 1

Stage	Observed Cues	Timestamp	Time (in s)	Points Awarded	Time in Stage 4	Percent accuracy	Percent relevant
3	Laceration on the right side	0.06	6	0		100	100
3	Blood is bright red	0:15	15	0		100	100
3	Jaundice	0:23	23	-1		0	0
3	Pale	0:26	26	1		100	100
3	Unaware of surroundings	0:31	31	0		80	100
3	Labored breathing	0:46	46	1		100	100
3	Dyspnea	0:49	49	1		100	100
3	Slow breathing	0:57	57	0		100	100
3	Tracheal deviation	1:16	76	1		100	100
3	Jugular vein distension	1:24	84	1		100	100
3	Low oxygenation	1:50	110	1		100	100
3	High heart rate	1:59	119	1		100	100
	MOVE TO STAGE 4	2:00	120		0		
4	0 oxygenation	2:00	120	1	0	100	100
4	Low heart rate	2:01	121	1	1	100	100
4	Eyes are more shut	2:05	125	1	5	60	50
4	More pale	2:10	130	1	10	80	50
4	Stopped breathing	2:19	139	0	19	100	100
4	Lower heart rate to 32 bpm	2:40	160	1	40	100	100
4	Tracheal deviation	2:47	167	1	47	100	100
4	Jugular looks the same	2:50	170	1	50	100	100
4	Unresponsive	2:56	176	1	56	100	100
4	Wound looks the same	3:10	190	0	70	100	0
4	**Diagnosis**: Pneumothorax	3:47	227	0.5	107		

Table 9b. Findings for participant 2

Stage	Observed Cues	Timestamp	Time (in s)	Points Awarded	Time in Stage 4	Percent accuracy	Percent relevant
3	Supine	0.15	15	0		100	50
3	Laceration on the right side under pectoralis	0:22	22	0		100	0
3	Oxygen is low	0:37	37	1		100	100
3	High HR	0:40	40	1		100	100
3	Disoriented	0:48	48	0		100	100
3	Confused	0:52	52	0		100	100
3	Doesn’t appear to be breathing in	1:05	65	0		100	100
3	Chest isn't moving much	1:10	70	0		100	50
	MOVE TO STAGE 4	2:00	120		0:00		
4	Stopped breathing	2:08	128	0	0:08	100	100
4	0 O2	2:18	138	1	0:18	100	100
4	HR down to 32 bpm	2:21	141	1	0:21	100	100
4	**Diagnosis**: Puncturing of right lung causing it to collapse	2:35	155	0.4	0:35		

Table 9c. Findings for participant 3

Stage	Observed Cues	Timestamp	Time (in s)	Points Awarded	Time in Stage 4	Percent accuracy	Percent relevant
3	Breathing slowly	0:05	5	0		100	100
3	Laceration under right peck	0:09	9	0		100	0
3	Tachycardic	0:14	14	1		100	100
3	Low O2	0:16	16	1		100	100
3	Lips are blue	0:18	18	1		100	100
3	Eyes open	0:20	20	0		100	0
3	No jugular vein distension	0:37	37	-1		0	100
3	No tracheal deviation	0:40	40	-1		0	100
3	Asymmetric breathing	0:54	54	1		100	100
	MOVE TO STAGE 4	1:24	84		0		
4	Low heart rate	1:25	85	1	1	100	100
4	Bradycardia	1:28	88	1	4	100	100
4	0 O2 saturation	1:33	93	1	9	100	100
4	Lips blue	1:37	97	1	13	100	100
4	Pale in face	1:38	98	1	14	100	100
4	Same laceration	1:43	103	0	19	100	100
4	Patient breathing very slowly	1:47	107	1	23	100	100
4	Chest not moving	2:01	121	0	37	100	100
4	Eyes closing	2:05	125	1	41	100	100
	Diagnosis: Hemorrhage from the cut bleeding into mediastinum causing pressure on the lungs inhibiting breathing	2:20	140	0	56		
	Treatment: Remove blood that's causing pressure	2:25	145	0	61		

Table 9d. Findings for participant 4

Stage	Observed Cues	Time stamp	Time (in s)	Points Awarded	Time in Stage 4	Percent accuracy	Percent relevant
3	Wound on right side	0:15	15	0		100	0
3	Slight bleeding	0:19	19	0		100	0
3	Elevated HR	0:26	26	1		100	100
3	Mouth open	0:36	36	0		100	0
3	Not breathing hard	0:41	41	1		100	100
3	Labored breathing	0:46	46	1		100	100
3	Eyes open	0:55	55	0		100	0
3	Stab wound/incision	1:17	77	0		100	0
3	Slightly pale	1:30	90	1		100	100
	MOVE TO STAGE 4	1:44	104		0		
4	Low HR	1:51	111	1	7	100	100
4	Unresponsive	1:54	114	0	10	100	100
4	Seems not to be breathing	1:55	115	0	11	100	100
4	Mouth open	2:36	156	0	52	100	0
4	Pale skin	2:46	166	1	62	100	100
	Diagnosis: Acute respiratory distress	2:06	126	0.2	22		

Table 9e. Findings for participant 5

Stage	Observed Cues	Timestamp	Time (in s)	Points Awarded	Time in Stage 4	Percent accuracy	Percent relevant
3	Male	0:06	6	0		100	0
3	Laceration on right peck	0:11	11	0		100	0
3	Blood is bright red	0:14	14	0		100	0
3	High HR	0:21	21	1		100	100
3	O2 is low at 85%	0:25	25	1		100	100
3	Breathing asymmetric	0:32	32	1		100	100
3	Breathing slow	0:26	26	0		100	100
3	Lips a touch blue	0:46	46	1		100	100
3	Disoriented	0:52	52	0		100	100
3	Mouth open	0:55	55	0		100	0
3	Trachea off-center	1:07	67	1		100	100
	MOVE TO STAGE 4	1:35	95		0		
4	HR dropped to 32	1:45	105	1	10	100	100
4	O2 is at 0	1:48	108	1	13	100	100
4	Breathing very slow	1:58	118	0	23	100	100
4	Laceration is the same	2:03	123	0	28	100	0
4	Blood is bright red	2:07	127	0	32	100	100
4	Trachea misaligned	2:11	131	1	36	100	100
4	Blue in face	2:16	136	1	41	100	100
4	Pale	2:21	141	1	46	100	100
4	Eyes closing	2:25	145	1	50	100	100
4	Mouth closing	2:28	148	0	53	100	0
4	JVD	2:43	163	1	68	100	100
4	Asymmetric chest	3:01	181	1	86	100	100
	Diagnosis: Pneumothorax	2:55	175	0.5	80		

At the end of the experiment, the participants were asked to answer debrief questions about the experience with the AR technology. Answers are shown in **Table 10.** Multiple participants thought it would be beneficial to be able to hear the patient breathing or be able to interact with the patient. Also, a few of the participants commented on the field of view being limited which hindered their ability to assess the entire body of the patient at one time.

Table 10. Responses to semi-structured interviews

	You diagnosed the patient with _________. Can you explain your reasoning for that diagnosis?	Were there any signs or symptoms that you wish were present in the augmented reality technology?	Were there any signs or symptoms that were present in the augmented reality technology that you felt were misleading?	For those in the experimental group: Did you feel the augmented reality training included: too much, not enough, or just the right amount of information to lead you to a diagnosis?	Do you have any other comments about your experience with the technology?
1	For pneumothorax: puncture wound on right side where lung would be, tracheal deviation, jugular vein distension, labored breathing, very pale, didn't seem oxygenated; For medial shift: couldn't see unevenness in medial spinal shift, but that usually accompanies pneumothorax	I wish I could hear him breathing. I wish I could get his temperature.	No	Just the right amount	Field of view is too small; When you get close to the patient, image becomes blurry; Pulse Ox and underwear kept coming off during training
2	Clearer movements of the chest indicating breathing on the test patient, but this patient didn't show movement of the chest at all while he was alive. O2 saturation was really low and HR was high indicating lung damage	Talk to him or hear him	No	N/A	Small field of view; Blurry when looking at it up close
3	Saw the cut on bottom of the right peck, initially saw asymmetric breathing, which caused an increase in HR and low blood oxygen. As it progressed, caused decrease in HR and lack of movement in chest	No	No	Just the right amount	Some type of movement; Patient talking; Hear him breathing (sound overall)
4	The textbook said symptoms include tachycardia and cardiac arrest. He had a wound in his right pectoral area which means lung has been damaged	Sound if possible; Hear breathing	No	N/A	No
5	The asymmetric breathing with only left side of chest elevating in addition of laceration and sharp decline in heart rate and blue skin tone indicated there was likely pneumothorax	Breathing sounds	No	Just the right amount	Manipulation of the patient (lift head/move arm); Ask him questions

DISCUSSION

Recent developments in the portability, flexibility, and affordability of AR technology make it a tool that can address the limitations of current medical training. As part of a multi-phase study, this pilot is focused on supporting simulation-based medical education through the use of AR technology. We aim to identify how this technology can enhance the detection and identification of subtle visual cues from the real world in the process of making accurate medical diagnoses. We randomly divided five participants into a control group, who received diagnostic training by reading a standard textbook, and an experimental group, who received diagnostic training by a combination of reading a standard textbook (same one as control group) and AR-based training. Participants were then presented with a tension pneumothorax scenario on the AR patient and were asked to list the symptoms they observed before making a final diagnosis. Investigators recorded both the total amount of time participants required to make a final diagnosis as well as the number of correctly identified symptoms, and the results of the control and experimental groups were statistically analyzed using a t-test to compare the groups. We found no significant difference between groups in the time needed to make a diagnosis. However, the AR-trained participants correctly identified a significantly greater number of symptoms (average: 12 symptoms) compared to the textbook-trained participants (average: 5 symptoms). These results indicate the potential of AR-based training to enhance detection and identification of cues in medical diagnostics.

The participants in the experimental group took more time to diagnose the AR patient than the participants in the control group. This finding is perhaps due to the AR training providing a more accurate and extensive list of symptoms on which the experimental participants were primed to focus. This phenomenon is therefore potentially a case of the speed-accuracy tradeoff, originally proposed by Paul Morris (Zhai, S., Kong, J., & Ren, X., 2004).

Although the experimental participants listed a significantly greater number of correct symptoms than the control group participants, neither group linked these reported symptoms to an accurate diagnosis. The correct diagnosis of the scenario used in the experiment was "tension pneumothorax," yet the most correct diagnostic response given was simply a "pneumothorax," which received half credit. This may have been from lack of repetition of the entire term "tension pneumothorax" in both the textbook and augmented reality-based training. In future research, we will intentionally emphasize the complete term "tension pneumothorax," and the difference between a tension pneumothorax and pneumothorax will be defined.

Three limitations were identified in this study. First, there were only 5 study participants, which decreases the accuracy of p-values. Second, all participants were current undergraduate students that had taken an anatomy course, but had never received prior training of clinical terminology, assessing a patient and proposing a concise diagnosis. In addition, these participants had not yet received the cardio-pulmonary training that is provided in January of the first year of medical school. Medical students do not participate in caring for patients in the work setting until the third year of medical school, so these participants were representative of first and second-year students with this characteristic, but third- and fourth-year medical students would be more advanced in their ability to assess patients. In the next phase of the study, 40 medical students will be tested, thereby increasing sample size and targeting participants with prior medical knowledge that can be applied to the AR scenario. Finally, the 120-second time limit on the first stage shown (Stage 3) proved to be too long, as several participants opted to progress early to Stage 4, allowing them to make a diagnosis at an earlier point. In the final study, the duration of Stage 3 will be reduced to a total of 60 seconds and participants will be required to use the full time.

The pilot study also highlighted several areas that need refinement before progressing to the larger study. Most importantly, the primary measurement needs to be altered. In this pilot study, our focus was time-to-diagnose. The primary measurement should instead be the number of correctly identified symptoms, as well as the accuracy of the diagnosis, which will shift our focus to the accuracy portion of the accuracy-speed tradeoff. Secondly, there were a few discrepancies (**Table 11)** between the MD-vetted rubric used to allot points and the participants' observations from the AR patient. The rubric will be revised to correspond to the AR display. If tension pneumothorax is being misrepresented with the AR technology, then we will need to work with Unveil LCC to correct these issues.

Table 11: Discrepancies between rubric and technology

What is currently in the rubric	**What participants were observing in tech**	**Stage**
Tachypnea/fast breathing	Slow breathing	3
N/A	Unaware of surroundings/disoriented/confused	3,4
Hypoxic	Chest not moving/stopped breathing	4
O2 saturation low	0 SpO_2	4
Blue lips/cyanosis	No blue fingertips or toes	3,4

This study also demonstrated the need to provide guidance in formulating a concise diagnosis. For example, we may ask a participant, “Please provide a specific, concise diagnosis for this patient, such as’ hemorrhage,” or we may ask for a “concise label” for the patient. These prompts will help clarify the participant’s thoughts and will aid in allotting points. An additional minor change is to allow participants to take notes on the textbook pages in an attempt to more accurately simulate a conventional method of study.

In the next phase of this study, we aim to determine if the use of AR technology improves symptom recognition and diagnostic accuracy. If our findings support the use of augmented reality, the findings could be used as a step in the direction of implementing this type of technology into medical schools.

References

Barrie, M., Socha, J. J., Mansour, L. A., & Patterson, E. S. (2019, September). Mixed Reality in Medical Education: A Narrative Literature Review. In *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* (Vol. 8, No. 1, pp. 28-32). Sage CA: Los Angeles, CA: SAGE Publications.

Barsom, E. Z., Graafland, M., & Schijven, M. P. (2016). Systematic review on the effectiveness of augmented reality applications in medical training. *Surgical Endoscopy, 30*(10), 4174-4183.

Crandall, B., Klein, G., Hoffman, R.R. (2006). Working minds: A practitioner's guide to cognitive task analysis. Cambridge, MA: MIT Press.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10), S70- 81.
Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge handbook of expertise and expert performance*, 683-703.

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review, 100*(3), 363.

Flach, J. M. & Voorhorst, F. (2016). *What matters? Putting common sense to work.* Dayton, OH: Wright State University Libraries.

Herth, F. J. F. (n.d.). Clinical presentation, diagnostic evaluation, and management of central airway obstruction in adults. Retrieved March 25, 2020, from https://www.uptodate.com/contents/clinical-presentation-diagnostic-evaluation-and-management-of-central-airway-obstruction-in-adults?search=airway obstruction&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1

Hoffman, R. & Militello, L.G. (2008). Perspective on Cognitive Task Analysis: Historical Origins and Modern Communities of Practice. New York: Taylor and Francis.

Hoffman, R. R., Ward, P., Feltovich, P. J., DiBello, L., Fiore, S. M., & Andrews, D. H. (2014). Accelerated expertise: Training for high proficiency in a complex world. New York: Psychology Press, Taylor & Francis Group.

Kellen R. Maicher, Laura Zimmerman, Bruce Wilcox, Beth Liston, Holly Cronau, Allison Macerollo, Lifeng Jin, Evan Jaffe, Michael White, Eric Fosler-Lussier, William Schuler, David P. Way & Douglas R. Danforth (2019) Using virtual standardized patients to accurately assess information gathering skills in medical students, Medical Teacher, 41:9, 1053-1059

Kirkpatrick, D.I. (1998). *Evaluating Training Programs: The four levels, 2nd edition*. San Francisco: Berrett-Koehler.

Kirkpatrick D. Techniques for evaluating training programs. In: Ely D Plomp T, eds. *Classic Writings on Instructional Technology*. Volume 1. Englewood, CO: Libraries Unlimited; 1996:119–42.

Klein, G., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics, 19, 462-472.

Klein, G., & Wolf, S. (1995, October). Decision-centered training. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 39, No. 20, pp. 1249-1252). Sage CA: Los Angeles, CA: SAGE Publications.

Kraiger, K. , Ford, J. K. , & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology, 78, 311-328.

Militello, L.G. & Hutton, R.J.B. (1998). Applied Cognitive Task Analysis (ACTA): A practitioner’s toolkit for understanding cognitive task demands. Ergonomics Special Issue: Task Analysis, 41 (11), 1618 -1641.

Militello, L. G., Patterson, E. S., Sushereba, C. E., & Wolf, S. P. (2018). *Principled Design of an Augmented Reality Trainer for Medics*. Technical Report [Contract No. W81XWH-17-C-0162]. Unveil LLC: Cincinnati United States.

Miller, G.E. (1990). *The assessment of clinical skills/competence/performance. Academic Medicine, 65*(9), S63-S67.

Patterson, E.S., Roth, E.M., Woods, D.D. (2010). Facets of complexity in situated work. In E. S. Patterson & J. Miller (Eds.) *Macrocognition metrics and scenarios: Design and evaluation for real-world teams.* Burlington, VT: Ashgate

Patterson, E. S., & Hoffman, R. R. (2012). Visualization framework of macrocognition functions. Cognition, Technology & Work, 14(3), 221-227.

Patterson, E.S., Roth, E.M., Woods, D.D.: Facets of complexity in situated work. In: Patterson, E.S., Miller, J. (eds.) Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams, pp. 1–10. Ashgate Publishing (2010)

Patterson & J. Miller (Eds.) *Macrocognition metrics and scenarios: Design and evaluation for real world teams.* Burlington, VT: Ashgate.

Patterson, M.D., Militello, L.G., Bunger, A., Taylor, R.G., Wheeler, D.S., Klein, G., Geis, G.L. (2016). Leveraging the Critical Decision Method to Develop Simulation- based Training for Early Recognition of Sepsis**,** *Journal of Cognitive Engineering and Decision Making, 10*(1), pp. 36-56.

Pliske, R. M., McCloskey, M. J., & Klein, G. (2001). Decision skills training: Facilitating learning from experience. Linking expertise and naturalistic decision making (pp 37-53). Psychology Press.

Rosen MA, Salas E, Wu TS, Silvestri S, Lazzara EH, Lyons R, Weaver SJ, King HB. Promoting teamwork: An event-based approach to simulation-based teamwork training for emergency medicine residents. Acad Emerg Med 2008b; 15: 1190–1198

Sauer, F., Khamene, A., & Vogt, S. (2010). *U.S. Patent No. 7,774,044*. Washington, DC: U.S. Patent and Trademark Office.

Shubert, J. (2019, March 14). Inhalation Injury. Retrieved March 25, 2020, from https://www.ncbi.nlm.nih.gov/books/NBK513261/

Zhai, S., Kong, J., & Ren, X. (2004). Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies*, *61*(6), 823–856.

Zhu, E., Hadadgar, A., Masiello, I., & Zary, N. (2014). Augmented reality in healthcare education: an integrative review. PeerJ, 2, e469.

Zsambok, C., & Klein, G. (1997). *Naturalistic Decision Making*. Hillsdale, NJ: Lawrence Erlbaum